UNIVERSITY OF CAMBRIDGE

MAMA/208, NST3AS/208, MAAS/208

MAT3

MATHEMATICAL TRIPOS Part III

Tuesday 10 June 2025 1:30 pm to 3:30 pm

PAPER 208

CONCENTRATION INEQUALITIES

Before you begin please read these instructions carefully

Candidates have TWO HOURS to complete the written examination.

Attempt no more than **THREE** questions.
There are **FOUR** questions in total.
The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury tag
Script paper
Rough paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 Concentration of information

For a random variable X with p.m.f. P, let $H(X) = \mathbb{E}(-\log P(X))$ denote its Shannon entropy. The *varentropy* of X is defined as

$$V_H(X) := \mathrm{Var}(-\log P(X)).$$

(a) State and prove Hoeffding's lemma for discrete random variables.

(b) Show that, if $X \sim \text{Bernoulli}(p)$, then $H(X) \leqslant 2\sqrt{p} + p$.

(c) For $i = 1, \ldots, n$, let $p_i \in (0, 1)$. Suppose $X_1, \ldots, X_n$ are independent Bernoulli(p_i) random variables such that X_i has probability mass function (p.m.f.) P_i, i.e. $\mathbb{P}(X_i = 1) = P_i(1) = p_i = 1 - P_i(0)$. Let $P = P_1 \otimes \cdots \otimes P_n$ denote the p.m.f. of $(X_1, \ldots, X_n)$.

Show that, if for some $\frac{1}{9} > \delta_1, \ldots, \delta_n > 0$, $H(X_i) \geqslant 3\sqrt{\delta_i}$ for every $i = 1, \ldots, n$, then for any $t \geqslant 0$,

$$\mathbb{P}\Big(\log P(X_1, \ldots, X_n) + H(X_1, \ldots, X_n) \leqslant -t\Big) \leqslant e^{-\frac{2t^2}{\sum_{i=1}^n (\log \delta_i)^2}}.$$

(d) Under the same assumptions as in part (c), show the following bound for the varentropy:

$$V_H(X_1, \ldots, X_n) \leqslant \frac{1}{4} \sum_{i=1}^{n} (\log \delta_i)^2.$$

[For parts (b),(c) and (d) you may use any results from the lectures, provided you state or quote them clearly.]

2 Poincaré and exponential concentration

Let X be a random vector on $\mathbb{R}^d$ with finite Poincaré constant, $C_P(X) < \infty$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable function, which is 1-Lipschitz (with respect to the Euclidean norm). Assume that $\mathbb{E}f(X) = 0$ and $F(\lambda) := \mathbb{E}e^{\lambda f(X)} < \infty$ for every $\lambda \in \mathbb{R}$.

[In all parts of this question, you may use any results from the lectures provided you state or quote them clearly.]

(a) Show that for every $m \geqslant 1$, $|\lambda| < \frac{2}{\sqrt{C_P(X)}}$,

$$F(\lambda) \leqslant \prod_{k=0}^{m-1} \left(\frac{1}{1 - \frac{\lambda^2 C_P(X)}{4^{k+1}}} \right)^{2^k} F\Big(\frac{\lambda}{2^m}\Big)^{2^m}.$$

[You may use without proof the fact that if f is differentiable and 1-Lipschitz, $\|\nabla f\| \leqslant 1$.]

(b) Show that for any λ,

$$F\Big(\frac{\lambda}{2^m}\Big)^{2^m} \to 1 \quad \text{as } m \to \infty$$

and for $\lambda^* = \frac{1}{\sqrt{C_P(X)}}$,

$$F(\lambda^*) \leqslant 3.$$

(c) Conclude that for every $t \geqslant 0$,

$$\mathbb{P}\big(|f(X)| \geqslant t\big) \leqslant 6e^{-\frac{t}{\sqrt{C_P(X)}}}.$$

(d) Let Y be another random variable on $\mathbb{R}$ having density $\frac{1}{2}e^{-|x|}, x \in \mathbb{R}$. Show that

$$C_P(Y) \leqslant 4.$$

[You may use without proof the identity $\mathbb{E}f(Y) - f(0) = \mathbb{E}\mathrm{sgn}(Y)f'(Y)$, where for $x \in \mathbb{R}$, $\mathrm{sgn}(x)$ denotes the sign of x.]

3 Probability distances

Let f, g be everywhere positive, continuously differentiable probability densities on $\mathbb{R}$. The relative entropy between f and g is

$$D(f\|g) := \int_{\mathbb{R}} f(x) \log \frac{f(x)}{g(x)} dx.$$

The *Fisher information distance* between f and g is

$$J(f\|g) := \int_{\mathbb{R}} f(x) \Big(\frac{f'(x)}{f(x)} - \frac{g'(x)}{g(x)} \Big)^2 dx.$$

The *Hellinger* distance between f and g is

$$d_H(f,g) := \left(\frac{1}{2} \int_{\mathbb{R}} \left(\sqrt{f(x)} - \sqrt{g(x)} \right)^2 dx \right)^{\frac{1}{2}}.$$

(a) State the Gaussian Poincaré and the Gaussian log-Sobolev inequalities.

(b) Let f be a probability density as above and $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$. Show that

$$D(f\|\phi) \leqslant \frac{1}{2} J(f\|\phi).$$

(c) Let f and ϕ as above. Show that

$$d_H^2(f,\phi) \leqslant \frac{1}{4} J(f\|\phi).$$

[You may use any results from the lectures provided you state or quote them clearly.]

4 Transportation and Couplings

(a) Let Z be an integer-valued random variable having distribution P such that $\mathbb{E}|Z|e^{\lambda Z} < \infty$ for all $\lambda \in \mathbb{R}$. Let $\psi_{Z-\mathbb{E}Z}(\lambda)$ denote the log-MGF of $Z - \mathbb{E}Z$ and let its Legendre transform be denoted by $\psi^*_{Z-\mathbb{E}Z}(t) = \sup_{\lambda\in\mathbb{R}}[\lambda t - \psi_{Z-\mathbb{E}Z}(\lambda)]$. Let

$$b := \sup\{\psi'_{Z-\mathbb{E}Z}(\lambda) : \lambda \in \mathbb{R}\}.$$

Prove that for all $0 < t < b$,

$$\psi^*_{Z-\mathbb{E}Z}(t) = \inf_Q\{D(Q\|P) : \mathbb{E}_Q Z - \mathbb{E}Z \geqslant t\}.$$

[You may use without proof the fact that $\psi'_{Z-\mathbb{E}Z}(\lambda)$ is continuous and the relative entropy is positive. Anything else that you use, you should prove.]

(b) Let $X \sim P$ and $Y \sim Q$ be random variables with values on a countable set Ω. Recall that the total variation distance is

$$d_{\mathrm{TV}}(P,Q) = \sup_{A\subset\Omega} |P(A) - Q(A)|.$$

Prove that

$$d_{\mathrm{TV}}(P,Q) \leqslant \inf_{\pi\in\Pi(P,Q)} \mathbb{P}_\pi(X \neq Y),$$

where $\Pi(P,Q)$ denotes the set of all couplings of P and Q.

(c) Recall that a random variable N follows the Poisson(ν) distribution if, for any $k \geqslant 0$,

$$P_N(k) = \frac{e^{-\nu}\nu^k}{k!}.$$

Let N be a Poisson(ν) random variable. For $i = 1, \ldots, n$, let $X_i \sim$ Bernoulli(p_i) be independent with $\sum_{i=1}^n p_i = \nu$ and let $S = \sum_{i=1}^n X_i$. Let P, Q denote the probability measures according to which N and S are distributed. Show that

$$d_{\mathrm{TV}}(P,Q) \leqslant \sum_{i=1}^n p_i^2.$$

[You may use without proof the fact that, if $N_i \sim$ Poisson(ν_i) are independent, then $\sum_{i=1}^n N_i \sim$Poisson($\sum_{i=1}^n \nu_i$) and any result from the lectures provided that you state it clearly.]

END OF PAPER